

Implementation of position assimilation for ARGO floats in a realistic Mediterranean Sea OPA model and twin experiment testing

V. Taillandier¹ and A. Griffa^{1,2}

¹CNR-ISMAR, La Spezia, Italy

²RSMAS, University of Miami, Florida, USA

Received: 23 February 2006 – Published in Ocean Sci. Discuss.: 18 Mai 2006

Revised: 8 September 2006 – Accepted: 13 November 2006 – Published: 22 November 2006

Abstract. In this paper, a Lagrangian assimilation method is presented and implemented in a realistic OPA OGCM with the goal of providing an assessment of the assimilation of realistic Argo float position data. We focus on an application in the Mediterranean Sea, where in the framework of the MFSTEP project an array of Argo floats have been deployed with parking depth at 350 m and sampling interval of 5 days. In order to quantitatively test the method, the “twin experiment” approach is followed and synthetic trajectories are considered. The method is first tested using “perfect” data, i.e. without shear drift errors and with relatively high coverage. Results show that the assimilation is effective, correcting the velocity field at the parking depth, as well as the velocity profiles and the geostrophically adjusted mass field. We then consider the impact of realistic datasets, which are spatially sparse and characterized by shear drift errors. Such data provide a limited global correction of the model state, but they efficiently act on the location, intensity and shape of the described mesoscale structures of the intermediate circulation.

1 Introduction

Lagrangian measurements have experienced a dramatic increase in recent years. The Argo program, consisting of a broad-scale global array of floats, is now one of the main components of the ocean observing system (<http://www.argo.ucsd.edu>). Argo floats are autonomous profiling floats that spend most of their life freely drifting at a given parking depth, and that surface at a regular time interval measuring temperature and salinity (TS) while ascending through the water column. When at the surface, they communicate information on position and TS profiles via satellite. Position

data provide information on the ocean current at the parking depth, while TS data characterize the water masses.

Since Argo floats provide Near Real Time (NRT) information, they are well suited to be used in operational systems for assimilation in Ocean General Circulation Models (OGCM's). Data on TS profiles are presently routinely assimilated in many operational systems (http://www.argo.ucsd.edu/FrUse_by_Operational.html) while position data are not included yet. On the other hand, float positions \mathbf{r} contain information on an important model state variable, the Eulerian velocity \mathbf{u} . The relationship between \mathbf{r} and \mathbf{u} is highly nonlinear, $d\mathbf{r}/dt = \mathbf{u}(\mathbf{r}, t)$, therefore the assimilation of \mathbf{r} poses significant challenges. The first theoretical works on position assimilation (Ishikawa et al., 1996) simply circumvented this problem using the quantity $\Delta\mathbf{r}/\Delta t$, where Δt is the interval between successive position information, as a proxy for \mathbf{u} . This approximation holds for very small Δt with respect to the typical Lagrangian time scale T_L (Molcard et al., 2003), which is of the order of 1–3 days at the surface and 5–10 days in the subsurface. In the case of Argo floats, since Δt and T_L are approximately of the same order, the approximation is clearly violated and appropriate assimilation methods that take into account the true nature of the observations have to be considered.

In the last few years, growing attention has been given and various methodologies proposed to assimilate position data. Different Lagrangian approaches have been investigated, going from Optimal Interpolation (OI) (Molcard et al., 2003, 2005; Özgökmen et al., 2003), Kalman filtering (Ide et al., 2002; Kuznetsov et al., 2003), to variational techniques (Kamachi and O'Brien, 1995; Nodet, 2006; Taillandier et al., 2006a). They have been tested mostly using simplified dynamical systems, such as point vortices (Ide et al., 2002), or academic model configurations (Kamachi and O'Brien, 1995; Molcard et al., 2005). Even though the results are very encouraging and show the high potential of assimilating Lagrangian data, there are still a number of issues to be

Correspondence to: V. Taillandier
(vincent@sp.ismar.cnr.it)

addressed regarding the actual application to Argo data assimilation in operational OGCM.

The first challenge to be addressed is the implementation of these assimilation methods in realistic circulation models. In particular, with respect to previous applications in idealized model configurations with a reduced number of isopycnal layers (Molcard et al., 2005), expected improvements should include complex geometry and topography, continuous stratification and TS mass variables in the model state. Other challenges are inherent to the nature and distribution of the Argo datasets. NRT position information is obtained when the floats reach the ocean surface and communicate via satellite. Such surface position data provide information on the subsurface drift (at the parking depth) contaminated by “shear errors” due to additional drifts experienced during ascent and descent motions (Park et al., 2005). The impact of such errors needs to be evaluated and possible techniques to account for it should be investigated. Also, Argo float positions are provided at intervals Δt of the order (or even greater) than T_L , so that successive \mathbf{r} along observed trajectories might be weakly correlated. Therefore the information content in terms of velocity might be degraded (Molcard et al., 2003). Finally, the realistic coverage of Argo floats in a given region of the world ocean is expected to be quite low to resolve the mesoscale features of the circulation. These different aspects have to be tested in order to verify the impact and potential of Argo position data if included in operational observing systems.

In this paper we present a Lagrangian assimilation method for realistic OGCM's. The various points indicated above are tested with the goal of providing a complete assessment of the assimilation of realistic Argo floats. The inverse method builds on previous works (Molcard et al., 2003, 2005; Taillandier et al., 2006a) and conceptually follows the same three main steps as in Molcard et al. (2005). First the position data \mathbf{r} are used to correct the velocity field \mathbf{u} at the parking depth z_p using the variational method developed in Taillandier et al. (2006a). Then the velocity correction is projected over the water column using statistical correlations while maintaining mass conservation. And finally the TS mass field is corrected in geostrophic balance with the corrected \mathbf{u} profiles using a variational approach (Talangrand and Courtier, 1987), based on the thermal wind and on the equation of state.

Here we consider a specific application to a realistic OPA model of the Mediterranean Sea. In the framework of the MFSTEP project (Pinardi et al., 2003; <http://www.bo.ingv.it/mfstep/>), a component of the world wide Argo project has been carried out in the Mediterranean Sea. A total of more than 20 floats have been deployed in the basin, characterized by a time interval Δt of approximately 5 days and by a parking depth z_p of 350 m (Poulain, 2005). In order to test the performances of the assimilation method in the “twin experiment” approach, synthetic floats with characteristics similar to the in-situ ones are released in a numerical circulation to be identified. The sensitivity of the Lagrangian analysis to

shear errors and low coverage is considered in a second time.

The paper is organized as follows. The method is presented in Sect. 2, while the set up for the numerical experiments is given in Sect. 3. The experiment results are presented in Sect. 4. Assimilation skills with realistic Argo type data are investigated in Sect. 5. Summary and discussion are provided in Sect. 6.

2 Methodological aspects

The analysed (estimated) circulation is obtained from float position information by correcting a prior (background) circulation in a sequential way. Such model state correction is provided every five days according to the cycling design of the Argo floats. Estimations of the velocity field \mathbf{u} and of the corresponding mass field (T, S) are detailed in the following sections 2.1 and 2.2 respectively for each 5 day sequence, while the practical implementation of the method is summarized in Sect. 2.3.

2.1 Estimation of the model velocity field

The velocity field at the parking depth z_p is estimated from Argo data using the variational method developed by Taillandier et al. (2006a). The method provides a bi-dimensional velocity increment $\Delta\mathbf{u}(z_p)$ for each 5 day sequence separating two successive Argo float positions. In this approach, $\mathbf{u}(z_p)$ is estimated minimizing a cost function which measures the distance between the observed float position at the end of the sequence and the position of a prior trajectory advected in the background velocity field. $\Delta\mathbf{u}(z_p)$ is therefore a time-independent correction (for the given sequence) of $\mathbf{u}(z_p)$ mapped around the prior trajectory. Note that shear drifts occurring on vertical float motions can be taken into account in the computation of the prior trajectory but only $\mathbf{u}(z_p)$ is estimated.

This variational method has been tested in Taillandier et al. (2006a), considering its performance by varying number of floats P and time sampling Δt . The results are positive, showing that the approach is effective and robust. Moreover, the method provides significant advantages with respect to the previous works of Molcard et al. (2003, 2005), since it expands the velocity correction all along trajectories and uses an appropriate covariance matrix. We notice that in Taillandier et al. (2006a) an important issue was discussed for the case of Δt of the order or bigger than T_L . This is of potential interest here given that T_L is expected to be approximately of the same order as Δt (i.e. five days). It is pointed out that when strong vortices are sampled by various floats, it can happen that successive positions “cross” each other because the vortex curvature is not resolved by the trajectory sampling. Such coherent structures are therefore not correctly estimated, since the obtained flow converges toward the crossing points. This problem, which is inherent to the

data, was illustrated in the case where the velocity field is reconstructed using data only, i.e. without prior circulation. In the experiments considered here, we expect that the use of prior model circulations, and consequently the use of additional information on the curvature, will help preventing this phenomenon. We will come back on this issue discussing the numerical results in Sect. 4.

In first approximation for regional and oceanic circulation scales, horizontal and vertical correlations of the motion can be separated using linear regression profiles (Oschlies and Willebrand, 1996). Note that horizontal velocity correlations are specified by the length of mesoscale structures described by the float trajectories (Taillandier et al., 2006a). As discussed in Molcard et al. (2005), the velocity increment $\Delta \mathbf{u}(z_p)$ obtained from float positions at z_p can be projected on the water column as

$$\Delta \mathbf{u}(z) = \mathbf{R}(z) \cdot \Delta \mathbf{u}(z_p), z \in [0, h] \quad (1)$$

where h is the depth of the basin and \mathbf{R} represents the vertical projection operator. $\mathbf{R}(z)$ is computed from seasonal and spatial averages (in quasi-homogeneous regions) of the linear regression coefficients linking each velocity component at depth z with velocity at z_p .

This projection of $\Delta \mathbf{u}(z_p)$ is funded on statistical means, without any constraint on model consistency. In order to introduce some dynamical constraints, we do enforce the assumption of basin volume conservation expressed by a non-divergent barotropic flow in the rigid lid formulation. To do that, the depth integrated velocity increment is first expressed from Eq. (1) as

$$\Delta \mathbf{U} = \Delta \mathbf{u}(z_p) \cdot \int_{[0,h]} \mathbf{R}(z) \cdot dz \quad (2)$$

Then its divergent part is removed while $\Delta \mathbf{U}$ is further recomputed from the diagnostic of its stream function. After that, the updated barotropic flow increment $\Delta \mathbf{U}$, now in model consistency, is used to redefine the velocity increment as

$$\Delta \mathbf{u}(z) = \mathbf{R}(z) \cdot \left[\int_{[0,h]} \mathbf{R}(z) \cdot dz \right]^{-1} \cdot \Delta \mathbf{U}, z \in [0, h] \quad (3)$$

This normalisation to the flat bottom case aims at modifying the amplitude of the velocity profile. It has been preferred to shifting the velocity profile. Notice that the two approaches are possible since the separation of single subsurface velocity information into barotropic and baroclinic contributions is underdetermined. The present approach allows to smooth horizontal discrepancies on shallow zones, especially along the land-sea interface at the parking depth, with a consistent velocity correction at the sea bottom.

2.2 Estimation of a dynamically consistent mass field

For mesoscale open sea applications like the present one, the isopycnal slopes can be adjusted, at least in first approximation, to the velocity increment $\Delta \mathbf{u}$ (expressed in

Eq. 3) in agreement with the thermal wind equation (Oschlies and Willebrand, 1996). The corresponding TS correction is searched on the sequence average since $\Delta \mathbf{u}$ is time-independent. So the mass field estimation is provided by the resolution of a static inverse problem formulated in the following variational approach.

Considering a prior model state (\mathbf{u}, T, S) , a variation $\delta \mathbf{u}$ of the velocity field infers a variation of its vertical shear $\delta \sigma$, defined as $\partial_z \delta \mathbf{u}$. This vertical structure is maintained by a density variation $\delta \rho$ in geostrophic equilibrium as

$$\delta \sigma = -g/\rho_o f \cdot \mathbf{k} \times \nabla \delta \rho \quad (4)$$

where g is the gravity, ρ_o the density of reference, f the Coriolis parameter, \mathbf{k} the upward unit vector and ∇ the horizontal gradient. The density variation $\delta \rho$ is linearly expressed with respect to the equation of state. In case of a best-fit polynomial formulation (e.g. the one of Jackett and McDougall, 1995), the first order perturbation of density around its prior value $\rho = \sum_{i,j} \alpha_{ij} \cdot T^i \cdot S^j$ is taken from

$$\delta \rho = \sum_{i,j} \alpha_{ij} \cdot T^{i-1} \cdot S^{j-1} \cdot (i \cdot S \cdot \delta T + j \cdot T \cdot \delta S) \quad (5)$$

where δT and δS are the respective variations on the prior temperature and salinity profiles. Notice that this perturbation equation can be obtained from a local linearization (for model values of T and S) of the equation of state used by the OGCM.

Here, we require that the shear variation $\Delta \sigma = \partial_z \Delta \mathbf{u}$ provided by the analysis of the float positions (Eq. (3)) is fitted to the geostrophic velocity shear variation defined by Eq. (4). This latter can be expressed using Eqs. (4–5) by $\delta \sigma = \mathbf{M} \cdot (\delta T, \delta S)$, where \mathbf{M} is a linear model. The inverse problem consists in finding the best combination $(\delta T, \delta S)$, in the neighbourhood of prior TS profiles, that minimises the cost function

$$J = (\Delta \sigma - \mathbf{M} \cdot (\delta T, \delta S))^T \cdot (\Delta \sigma - \mathbf{M} \cdot (\delta T, \delta S)) \quad (6)$$

where T is the vector transpose. In numerical practice, J is minimised using the steepest descent algorithm M1QN3 (Gilbert and Lemaréchal, 1989), in the direction of the gradient

$$\nabla J = -\mathbf{B} \cdot \mathbf{M}^T \cdot (\Delta \sigma - \mathbf{M} \cdot (\delta T, \delta S)) \quad (7)$$

where the linear model \mathbf{M}^T is governed by the adjoint equations of Eqs. (4–5), and \mathbf{B} is classically defined by background error covariances. For the construction of \mathbf{B} , cross-covariances between temperature and salinity are not considered as the balance on these two variables is explicitly constrained by the equation of state (Eq. (5)). So only the univariate components are required and described by the field variance. Notice that they intrinsically define the vertical weighting for the contribution of the velocity shear fitting. In numerical practice, seasonal and space averages of the standard deviation profiles for TS define the diagonal matrix \mathbf{B} .

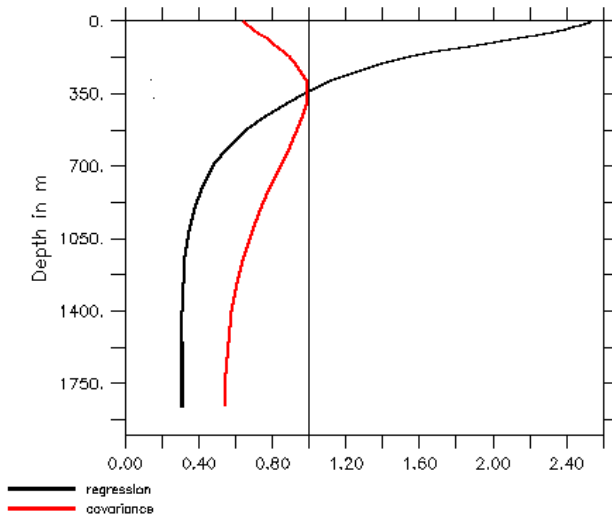


Fig. 1. Correlation and regression profiles of velocity components between depths z and $z_p=350$ m.

2.3 Assimilation procedures

Estimates of the prior circulation for each 5 day sequence have been implemented in two different assimilation procedures, referred to as “state of art” and “operational” respectively. Notice that each sequence initialization is stable since the corrected fields belong to the admissible model solutions. Both procedures have been tested in the twin experiments presented in the following.

In the procedure “state of art”, variations of Argo float positions during each cycle/sequence are used to estimate the velocity and mass fields at the initial time of the assimilation sequence. The model is then run forward again for the considered sequence using the corrected initial conditions, so that prior model state and trajectories are updated with respect to the initial state estimate. This procedure is characterised by a high computational cost as the CPU time of a forward model run is doubled, in addition to the estimation of the velocity and mass corrections. Note that CPU time requirements for a model state analysis is very small compared to a forward run.

In the procedure “operational”, variations of float positions during each cycle/sequence are used to estimate the velocity and mass fields at the final time of the assimilation sequence. Prior model state computed on the sequence determines the model update on the next sequence, and the prior float trajectories are not updated. This procedure is characterised by an equivalent computational cost as for a single forward run, always in addition to velocity and mass estimation.

3 Experimental set up

3.1 Modelling the Mediterranean circulation

The numerical model used here is an extended version of the primitive equation model OPA with rigid lid (Madec et al., 1998), configured in the Mediterranean Sea. For the development and testing of the assimilation procedure, which is computationally demanding, a horizontal resolution at $1/8$ degree has been considered, while the statistical prior information on \mathbf{R} and \mathbf{B} have been computed using a more refined resolution at $1/16$ degree. For both configurations a vertical discretization with 43 geopotential levels is considered. The model is initialised from rest by climatological hydrological conditions from MEDATLAS. It is forced by daily winds and heat fluxes from the ECMWF atmospheric forcing fields during the period 1987–2004. The spin-up has been performed repeating the cycle three times. Characteristics of the general circulation and of the transports at several straits are in good agreement with recent estimates from in-situ observations (Béranger et al., 2005).

We concentrate on a case study in the North Western Mediterranean Sea, during the winter season. This region has been chosen since it is a complex and interesting area with a vigorous mesoscale field even at 350 m. Moreover, a good coverage of in-situ Argo floats has been maintained in this region during the MFSTEP project (see the 4 float trajectories on <http://poseidon.ogs.trieste.it/WP4/product.html>). As a first step, the statistical quantities have been computed from model outputs during 2000–2004. Horizontal motion scales, already described in Taillandier et al. (2006a) are characterized by typical mesoscale length scales of the order of 20–30 km, and associated Eulerian time scales of the order of 20–30 days. Vertical correlations are computed from the velocity fields to define \mathbf{R} (Eq. (1)), and from the active tracer fields TS to define \mathbf{B} (Eq. (7)). These correlations are computed independently for each velocity or tracer component, and averaged over the north western basin, to get winter mean representative profiles. As underlined in Sect. 2.2, only the variance profiles for tracer fields are involved in the construction of \mathbf{B} . In the region of interest, maximum variance for temperature reaches 0.50°C in the first 100 m, followed by a quick decrease to 0.15°C in the intermediate layer. For salinity, the decrease from the maximum value of 0.12 PSU at surface is quasi linear, to reach 0.05 PSU at the parking depth.

Regression and correlation profiles for velocity are represented in Fig. 1. They are very similar for the zonal and meridional components. The velocity variations on the intermediate layer, between 150 m and 750 m, appear well correlated with velocity variations at the parking depth. Instead in the surface layer and in the deep layer, correlation values drop under 0.8. The linear regression coefficients decrease from 2.5 at sea surface until 0.4 at sea bottom. The profile can be described by four straight lines: a surface segment

between 0 m and 200 m associated to a sharp slope, an intermediate segment extending to 600 m, a deep segment extending to 1000 m associated to a slight slope, and a bottom segment after 1000 m without slope. Considering Eq. (1), the first order derivative of the regression profile is related to the relative intensity of the vertical shear variations. As a consequence, the correction of the baroclinic velocity (Sect. 2.2) will be limited to the first 1000 m, and it is expected to be characterized by decreasing amplitude from the sea surface, as in three distinct layers quite well correlated.

3.2 Numerical experiments

The sensitivity and robustness of the assimilating system has been tested with a total of seventeen simulation experiments. A first set of experiments (Sect. 4) is directly finalized at testing the methodology and the two different implementations of Sect. 2.3. It considers “ideal” data, i.e. data without observational errors and with significant spatial coverage. On the other hand, a second set of experiments (Sect. 5) is aimed at investigating the performance in realistic conditions, i.e. with reduced coverage, for Argo float applications. Moreover, surface position data are used instead of positions at the parking depth. In this case, observational information is contaminated by shear errors related to additional drifts experienced during ascent and descent motions. Notice that real in-situ Argo floats have another source of observational error due to the surface drift experienced before (after) the first (last) satellite fix. In this study, the surface drift error is not considered for simplicity, since it depends on the specific details of Argo communication. Our observational errors should then be considered as lower bounds with respect to in-situ errors. The numerical protocol of experimentation follows the classical twin experiment approach, and can be summarized as follows. Three different simulations are performed for each experiment using the circulation model (Sect. 3.1), all with the same forcing functions corresponding to the month of March 1999. The first run is correctly initialized from the model state corresponding to March 1 1999, and it represents the “truth” circulation \mathbf{u}^{tru} . Synthetic floats are launched in the truth run and advected by the numerical velocity field. The second run is initialized from a “wrong” initial condition, corresponding to the model state of 1 March 2000, to represent our incomplete knowledge of the true state of the ocean. The position data extracted from the truth circulation are assimilated in this run, providing the “estimated” field \mathbf{u}^{est} . Finally, the third run provides a “background” circulation \mathbf{u}^{bck} initialized with the same wrong initial condition as the assimilation run, but without data assimilation. It provides a reference simulation, where the circulation of March 1999 evolves freely from the wrong initial conditions. The success of the assimilation is quantified in terms of convergence of \mathbf{u}^{est} toward \mathbf{u}^{tru} . In other words a successful assimilation is expected to correct the wrong initial conditions, driving the ocean state toward the truth.

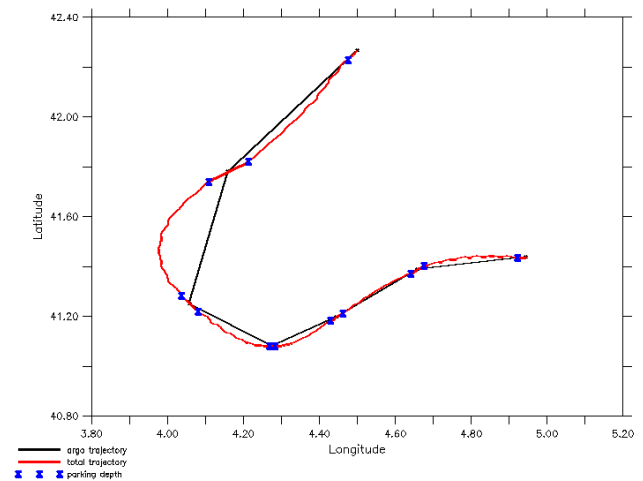


Fig. 2. Details of a simulated Argo type trajectory. The high resolution trajectory is shown, with superimposed the trajectory sampled at $\Delta t=5$ days, as observed at the surface. Crosses indicate positions where the simulated float reaches or leaves its parking depth.

The numerical trajectories are computed following the same cycle as in-situ Argo floats, using a fourth order Runge-Kutta scheme for advection. The cumulated duration of vertical motion is fixed to half a day, while the residence time in subsurface is equal to 4.5 days. In our numerical experiments two sets of data have been extracted from the numerical trajectories: the “perfect data”, containing initial and final positions of each subsurface drift, that are utilized in the idealized experiments (Sect. 4), and the surface data, that take into account the shear drift and that are used in the realistic experiments (Sect. 5). For in-situ floats, only surface positions are available, provided in near-real time via satellite. In Fig. 2, an example of the difference between the “perfect” data (marked by crosses) and “realistic” data observed at the surface is shown along a single numerical trajectory. A total of 182 simulated Argo floats have been advected in the truth circulation during six cycles (Fig. 3). In the assimilation experiments, the influence of data coverage is studied varying the number of Argo type trajectories P equal to 182, 42, 9, 4, 3 extracted from the original set. For all the cases, simulated floats are released on the basin according to homogeneous spatial distributions. This disables possible strategies of float deployment to a priori improve observational information (e.g. Toner et al., 2001; Molcard et al., 2006). As shown in Fig. 3, the Lagrangian description of the circulation shows an intense boundary current (the so-called Northern Current) characterized by meanders and recirculation cells. Given the long Eulerian time scale (20–30 days), the trajectories are indicative of mesoscale structures, but they quickly tend to get sparse when P decreases. On the other hand, the difference between “perfect” and “realistic” data, shown in

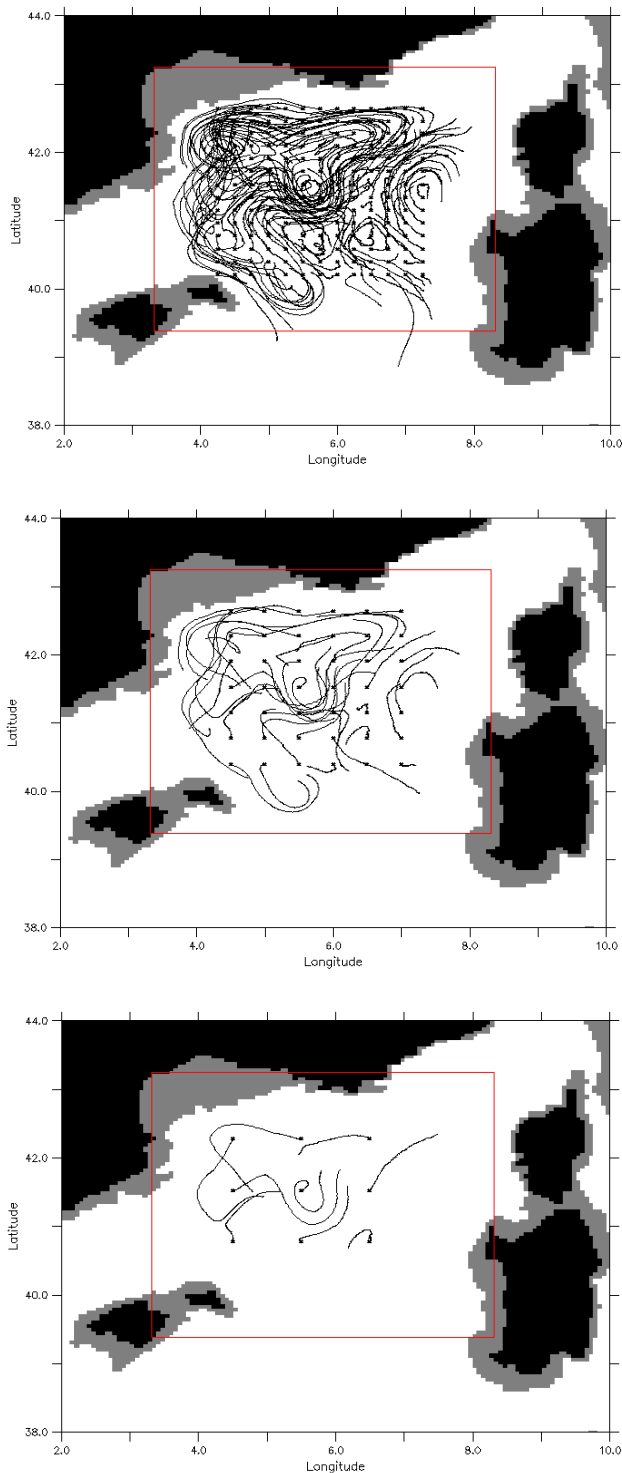


Fig. 3. Lagrangian description of the truth circulation \mathbf{u}^{tru} during the month of March 1999. Three coverages with different number of trajectories P are represented: $P=182$ (a), $P=42$ (b), $P=9$ (c). The locations of float deployment are indicated by a cross. The computational domain is bounded inside the box.

Fig. 2 has been assessed quantitatively. The ratio between the mean shear drift and the mean drift at the parking depth reaches 19% considering the 182 simulated trajectories.

The assimilation system is quantitatively assessed using a measure of the adjustment of \mathbf{u}^{est} to \mathbf{u}^{tru} . It is expressed by the non dimensional quantity

$$E(z, t) = \frac{\|\mathbf{u}^{\text{tru}}(z, t) - \mathbf{u}^{\text{est}}(z, t)\|^2}{\|\mathbf{u}^{\text{tru}}(z, t) - \mathbf{u}^{\text{bck}}(z, t)\|^2} \quad (8)$$

where $\|\mathbf{u}\|^2 = \mathbf{u}^T \cdot \mathbf{u}$. The normalization term in denominator is used to reduce the dependency of the results from the specific realization, i.e. from the time evolution of the background simulation. This measure is integrated over the computational region shown in Fig. 3. It can be applied at a given depth z or at a given sequence t . Notice that the time resolution is given by the duration of the sequence, so E values involve sequence averages for \mathbf{u} .

4 Results of experiments with ideal data

The methodology (Sect. 2) is tested using perfect data (i.e. without shear error) and considering the effects of decreasing coverage in a range of relatively high number of floats P equal to 182, 42, 9. The two different procedures discussed in Sect. 2.3 are compared. The results are assessed first in terms of velocity estimation at the parking depth and then in terms of velocity and mass estimation over the whole water column.

4.1 Velocity estimation at the parking depth

The first assessment focuses on the analysis of the circulation at the parking depth z_p . Results for P equal to 182, 42, 9 are shown in Fig. 4 in terms of time evolution of the adjustment error $E(t)$ (defined in Eq. 8) during the one month assimilation period using the two procedures. For both of them, the adjustment of the velocity field at 350 m is increasingly better with data coverage, as expected. The mean slope of the error decrease varies from 20% with low coverage (P equal to 9) to 35% with high coverage (P equal to 182). When considering the sequence by sequence evolution of $E(t)$, it appears that the major adjustment occurs during the first two analyses. After that, $E(t)$ continues to decrease in average, but oscillations can occur especially for P equal to 42 and 9, with slight increases followed by marked decreases. This can be explained considering the following points of view. In a fix point approximation (in which information locations do not change), the time scale characterising the assimilation adjustment is expected to coincide with the time scales of the circulation structures, i.e. with the time over which the information is correlated. With high coverage (P equal to 182), the fix point approximation can be assumed, and the obtained adjustment appears characterised by a constant decrease at the Eulerian time scale (~ 20 days). With low coverage instead, and given the relatively long Eulerian time scales, the velocity field can be assumed time-independent with respect to the

scale of motion of the floats. In this frozen field approximation, local and fast adjustments to the truth circulation are provided at the information location and at the Lagrangian time scale (order of 3–5 days). Successive adjustments are expected to occur when the floats enter and sample new circulation structures, providing new information. This appears to be the case for P equal to 9 or 42, where the decrease of $E(t)$ is characterised by secondary gaps due to the renewal of data information.

A comparison of the two assimilation procedures can now be performed. In the procedure “operational” (Fig. 4b), the model state correction is operated at the end of each sequence (Sect. 2.3), so that the effect of the model state correction for sequence (t) appears at sequence $(t+1)$. So this procedure only benefits of five analyses during the assimilation period, instead of six as for the procedure “state of art” (Fig. 4a). In terms of adjustment, the evolution of $E(t)$ appears similar in both procedures but translated of one sequence. So the adjustments E obtained at the final time with the procedure “operational”, 70% for P equal to 9, 40% for P equal to 42, 20% for P equal to 182, are approximately the same as the ones obtained a sequence before with the procedure “state of art”.

An example of qualitative comparison between estimated and truth circulations at the parking depth is drawn for the representative data coverage of P equal to 42 in Fig. 5. These 5 day-averaged circulations are mapped during the last sequence of the assimilation period. The results are similar for the two procedures, even though slight differences still remain on intensity and shape between the two estimated circulations. In both cases the assimilation appears to efficiently correct the field, and the estimates tend to converge toward the truth field. The Northern Current confined along the coast in the background circulation appears wider in the estimated circulations. Its meanders and recirculation features present in the truth circulation appear well estimated. Both intense and weak mesoscale structures are replaced from their background location with respect to the truth circulation.

We notice that in all the performed experiments, there is no evidence of problems in the velocity estimation when $\Delta t \geq T_L$ as discussed in Taillandier et al. (2006a) (and mentioned in Sect. 2.1) did not emerge. No unrealistic convergence in the estimated velocity field has been found. This might be partially due to the fact that the case in Taillandier et al. (2006a) was extreme in terms of intense sampling of the core of a very energetic vortex, even though also in the present experiments strong recirculation features and vortices are present and sampled at various rates, as shown in Fig. 3. More conceptually, we think that the main reason for the difference is that in Taillandier et al. (2006a), the velocity field was reconstructed based on the float data only, while in the present experiments the data are assimilated in a model. The model provides a priori information on the velocity structure, so that the information on the curvature can be corrected avoiding the emergence of unrealistic velocity

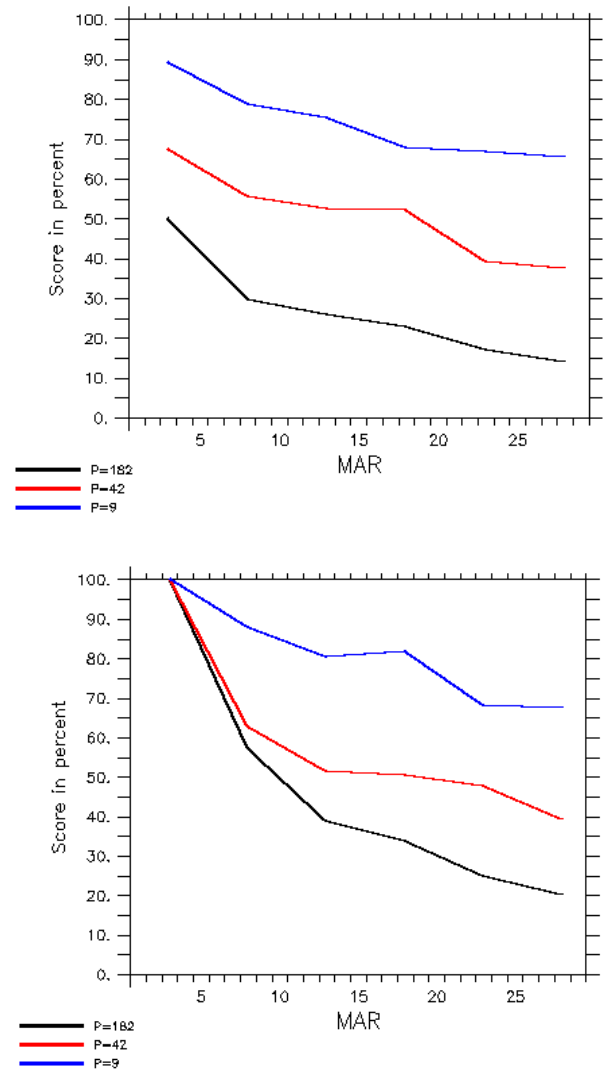


Fig. 4. Evolution of the adjustment error $E(t)$ (in %) at the parking depth computed from perfect data with three coverages $P=182$, 42, 9, using the assimilation procedure “state of art” (a) and “operational” (b).

convergence and a correct estimate can be obtained.

4.2 Velocity estimation on the water column and impact of TS correction

The effect of the velocity estimation at the parking depth is now investigated on the other layers. The quality of the model state analyses from float positions depends on the velocity adjustment at depth, which is given directly by the shape of linear regression profiles (see Eq. (1)) and indirectly by the mass field adjustment. This latter mechanism can either be forced through the estimation of temperature and salinity (Sect. 2.2), or lead by a natural adjustment of the mass field onto the estimated velocity field. In order to

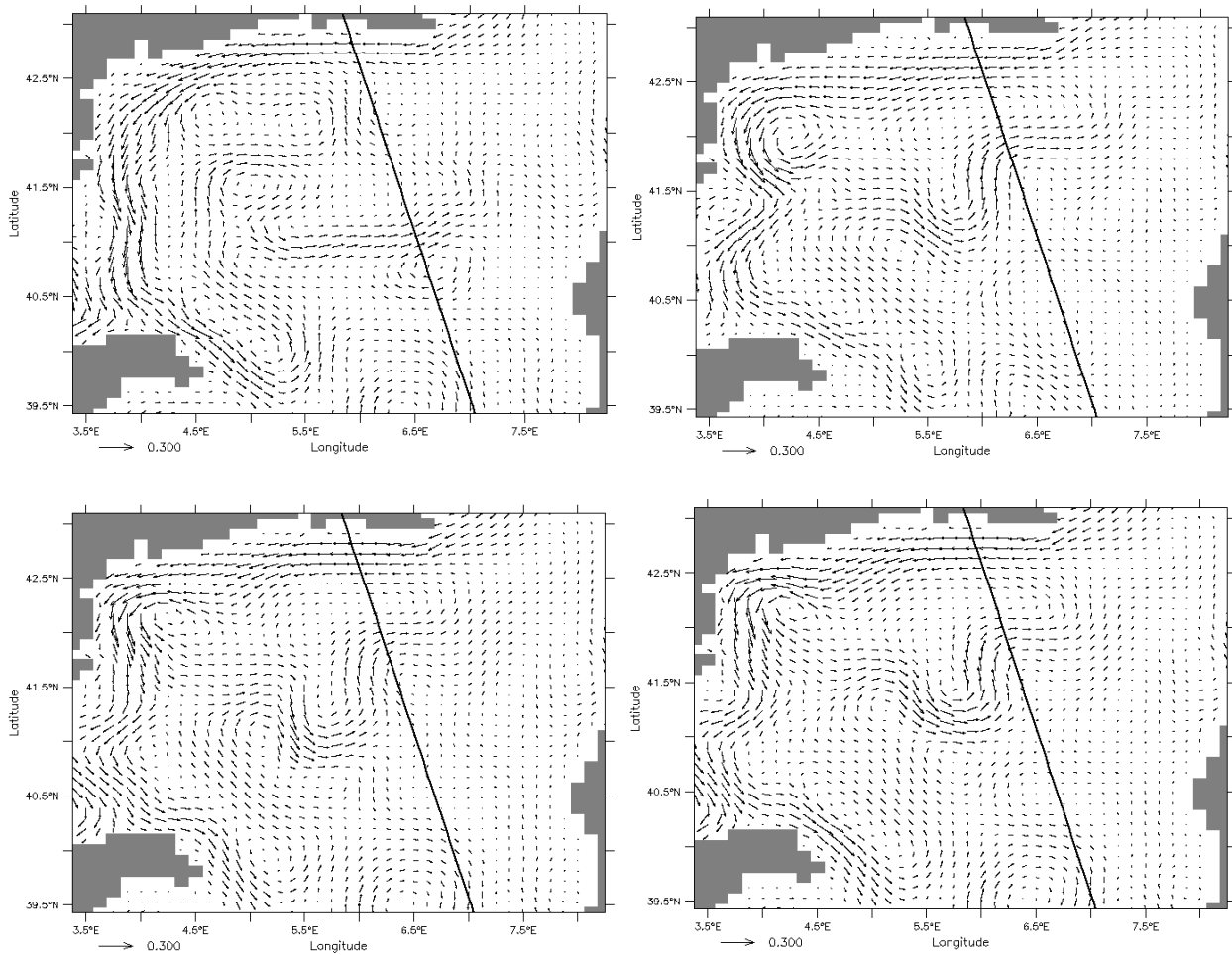


Fig. 5. Velocity fields at the parking depth at the end of the assimilation period. The background circulation is represented in panel (a), the truth circulation in panel (b), and the estimated circulation with the procedure “state of art” (“operational”) in panels (c) (d) using 42 trajectories. The hydrologic section is drawn as a straight line.

evaluate the impact of TS estimation, both approaches are tested when assimilating the three data sets P equal to 182, 42, 9. The corresponding adjustment errors $E(z)$ (given in Eq. (8) at the end of the assimilation period) are represented in Fig. 6. It appears that the profiles $E(z)$ are improved when obtained with TS estimation. This is particularly the case when the procedure “operational” is used. On the other hand, there is only a slight improvement when the procedure “state of art” is used. So the relaxation of the mass field is effectively operated during the sequence reanalysis, i.e. while it is set to the velocity estimate on the same sequence (Sect. 2.3). These adjustment errors appear to decrease from the sea surface to a minimum value close to the parking depth, increasing then with depth. This is particularly accentuated with high data coverage (P equal to 182), while the adjustments tend to be homogeneous on the water column with low data coverage (P equal to 9). This general feature can be explained considering the velocity correlations (Fig. 1) impli-

cated in the vertical velocity projection (Eq. 1). As described in Sect. 3.1, the amplitudes of the velocity variations at the sea surface are at least twice than those in the intermediate layer, while surface and intermediate currents are weakly correlated. So adjustment errors obviously increase in the surface layer. On the other hand, $E(z)$ values on the deep layer are mainly governed by the barotropic flow adjustment in which the baroclinic structure of the model state analyses is involved. So the quality of velocity estimates in the deep layer is degraded because of the estimation of the baroclinic structure, through the specification of vertical velocity shear, which is poor at depth (Sect. 3.1). Moreover, this degradation of $E(z)$ in the deep layers is accentuated when the adjustment of the model mass field is not effective. As shown for the “operational” procedure (Fig. 6 left panels), adjustment errors of deep currents are increased in absence of mass field estimation. A quantitative comparison of the effectiveness of the two assimilation procedures versus z can be performed

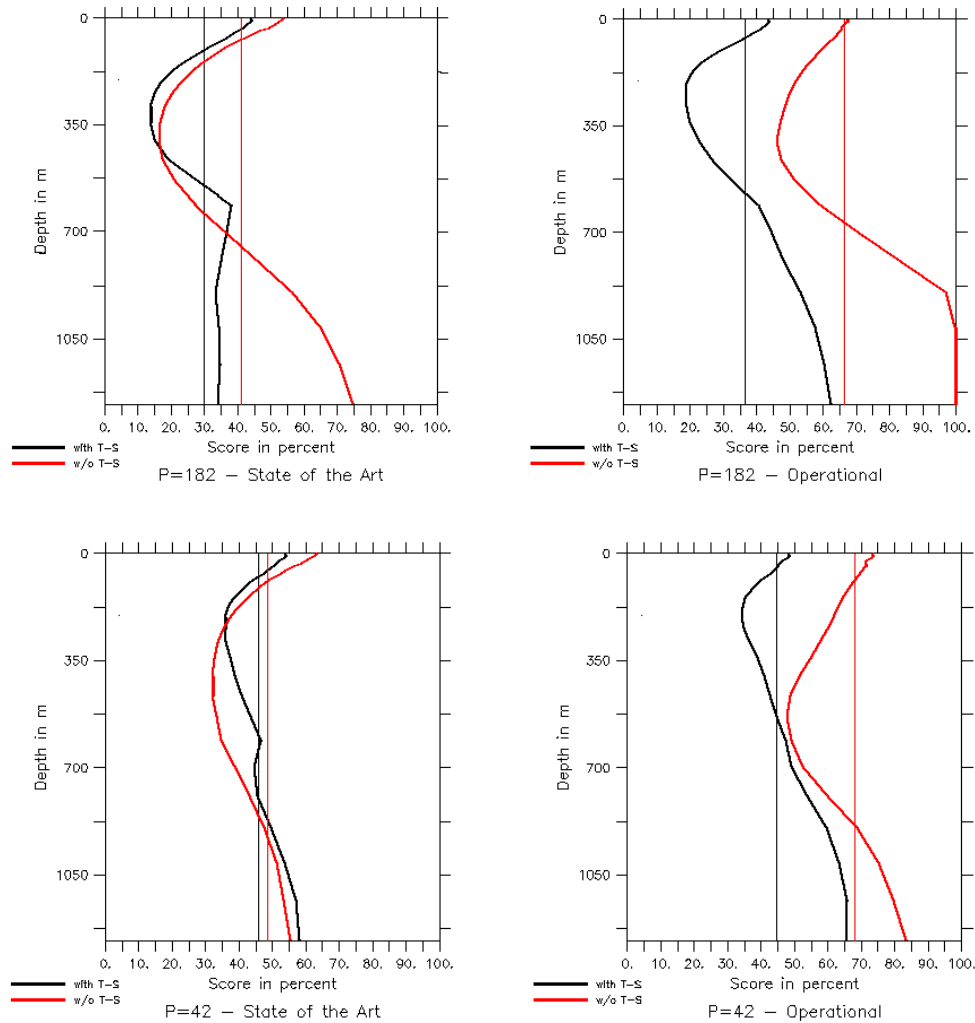


Fig. 6. Profiles of $E(z)$ at the end of the assimilation period for $P = 182$ (top), 42 (middle), 9 (bottom), for the procedure “state of art” (left) and “operational” (right). Adjustment errors of the vertical mean velocity are indicated in straight lines.

comparing the results in the left and right panels of Fig. 6. As it can be seen, results at the parking depth (discussed in Sect. 4.1) appear to hold also for an extended layer between 150 m and 500 m. For both procedures, the adjustment errors increase at the sea surface and in the deep layer. With low data coverage, this growth is limited and the results of the two procedures are similar. On the other hand, when the coverage is high (P equal to 182, 42), the growth of $E(z)$ at depth is limited ($\sim 20\%$) with the “state of art” procedure while is higher in the “operational” procedure ($\sim 35\%$). Notice that adjustments of vertical mean currents are of same order for both procedures. In order to better visualize the TS correction, a hydrologic section is extracted along the NW-SE transect shown in Fig. 5. The background, truth and estimated mass fields given at the end of the assimilation period are presented in Fig. 7 for the case of P equal to 42. Note that estimates of the mass field and of the horizontal velocity (Fig. 5) are ob-

tained from the same data set. Then each mesoscale structure crossed by the section appears geostrophically maintained in the intermediate layer by the isohaline and isothermal slopes. The estimated salinity and temperature fields appear similar to the truth, especially in the upper layer, even when the iso-lines are surfacing. In the intermediate layer, on the other hand, differences with the truth structures appear, especially below 400 m. As mentioned in Sect. 3.1, the correction of baroclinic structures sees its amplitude decreasing with depth with respect to the shape of velocity correlation profiles. So these differences on the mass field are not surprising given that the truth and background fields still strongly differ.

The misfit between the truth and background salinity field along the same section is shown in Fig. 8a, and compared with the misfit between estimated and background fields for two data coverages P equal to 42 (Fig. 8b) and P equal to 9 (Fig. 8c). For successful assimilation, the estimated misfit

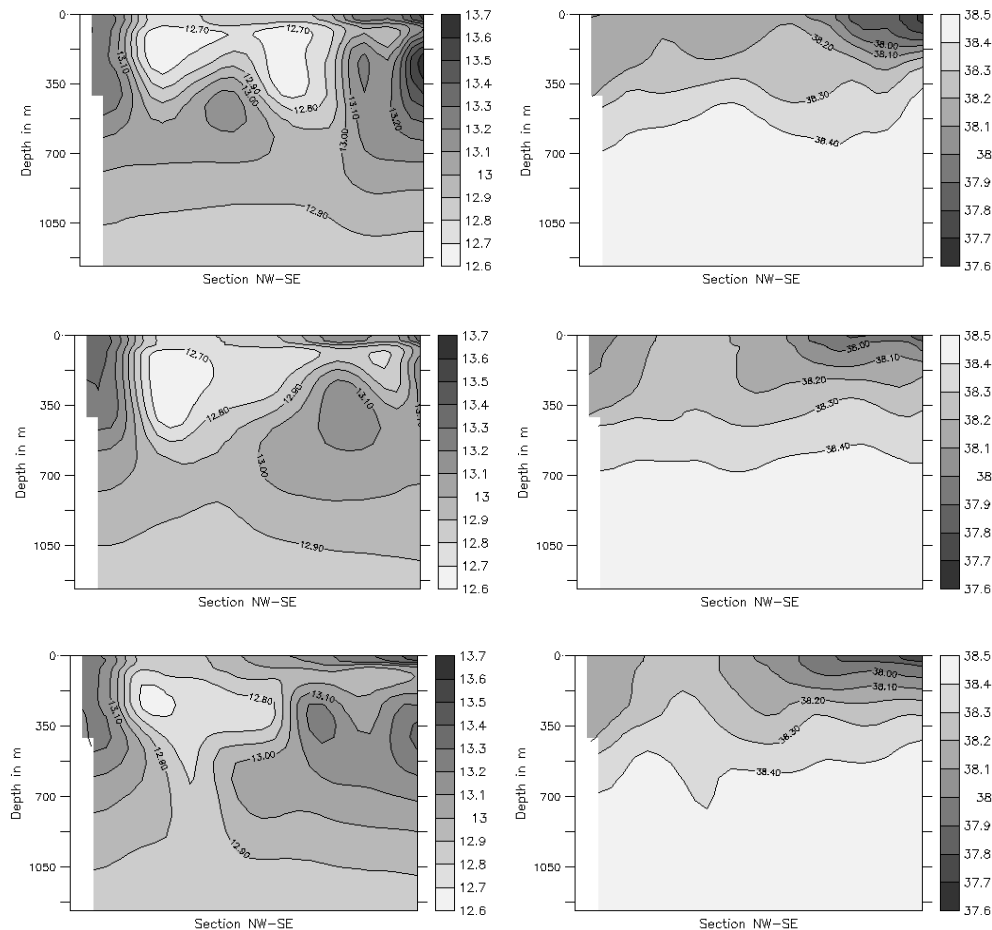


Fig. 7. Temperature in °C (left) and salinity in PSU (right) along the hydrologicSect. (see Fig. 5) at the end of the assimilation period for the background (top), truth (middle) and estimate with $P=42$ (bottom).

is expected to converge toward the truth misfit. In the upper layer the structures appear well reproduced for both coverages, even though for P equal to 9 the intensity is underestimated. Moreover, each surface misfit appears balanced in the subsurface, which contributes to the differences obtained in the intermediate layer, Fig. 7. Such mechanism is accentuated at low data coverage, and it can be related to the weak adjustment of baroclinic structures obtained with P equal to 9 (as shown Fig. 6, lower panels). In Fig. 8c, the hydrological patterns appear only estimated locally. At each of these locations, the water column sees its stratification translated vertically to generate isopycnal slopes in geostrophic equilibrium with the velocity correction. In consequence, the resulting density anomaly at the surface layer is balanced at depth.

In summary, the results show that the assimilation is effective. Velocity estimates at the parking depth converge toward the truth at increasing coverage. At depth the adjustment is effective in a large intermediate layer (from 100 m to 700 m) for high coverage, while it is weaker and more homogeneous for lower coverage. The TS correction is in good agreement

with the truth mass field, especially in the surface and intermediate layers, even for low coverage. Notice that improved adjustments in the surface layer would also be consecutive to the positive effect of unperturbed forcing fields (heat and momentum fluxes). The “operational” procedure, with a reduced CPU cost, appears to perform satisfactorily, provided that the active TS correction (Sect. 2.2) is performed.

5 Assimilation skills with realistic Argo type data

In this set of simulations, data with realistic characteristics with respect to Argo floats are considered. On the basis of previous results in Sect. 4, only the operational procedure is considered.

Results presented in the previous Sect. 4 were obtained using perfect positions, extracted at the parking depth. Here, instead, realistic surface data positions are considered, accounting for the “observational error” induced by shear drift. The methodology used for assimilation is the same as for perfect data, except that now the background trajectories used to

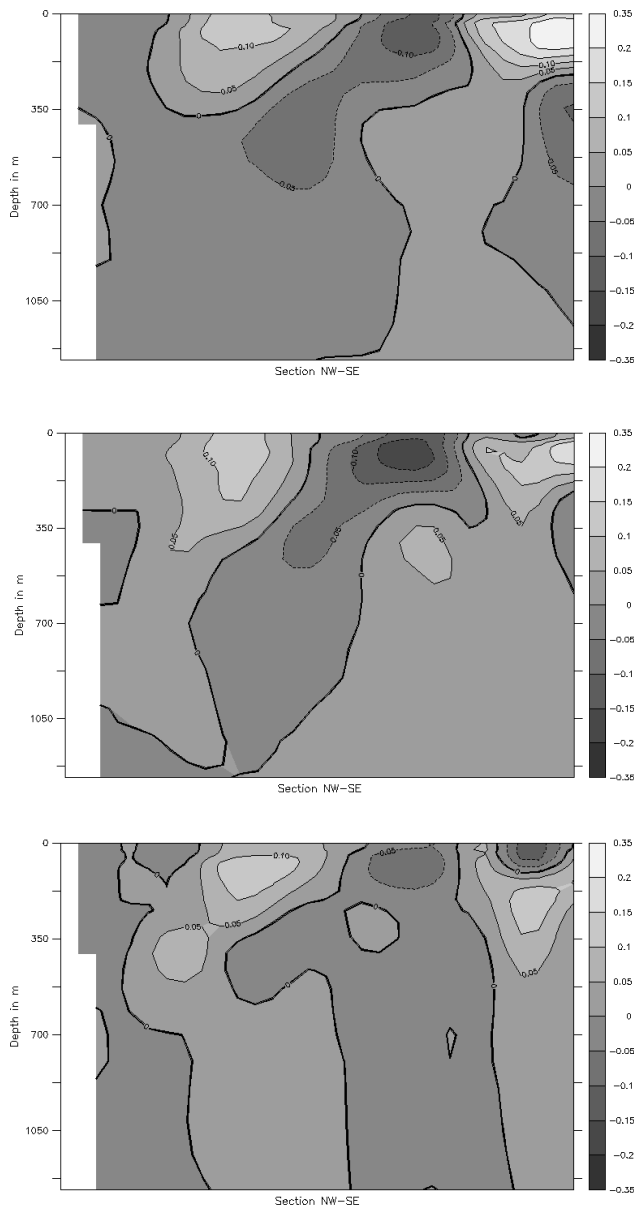


Fig. 8. Salinity misfits (in PSU) along the hydrologic section (see Fig. 5) at the end of the assimilation period: truth minus background (a); estimate minus background with $P=42$ (b) and with $P=9$ (c).

estimate the velocity at the parking depth (Sect. 2.1) contain ascent and descent motions at each sequence. Accounting for vertical motions is expected to provide prior information on the shear error, in the sense that the effects of shear are effectively inserted in the cost function (defined in Taillandier et al., 2006a) similarly to the classical specification of observation error covariances (Talagrand, 1999).

Adjustment errors $E(t)$ computed at the parking depth in presence of this observational error are presented in Fig. 9a for P equal to 182, 42, 9. They can be compared with the

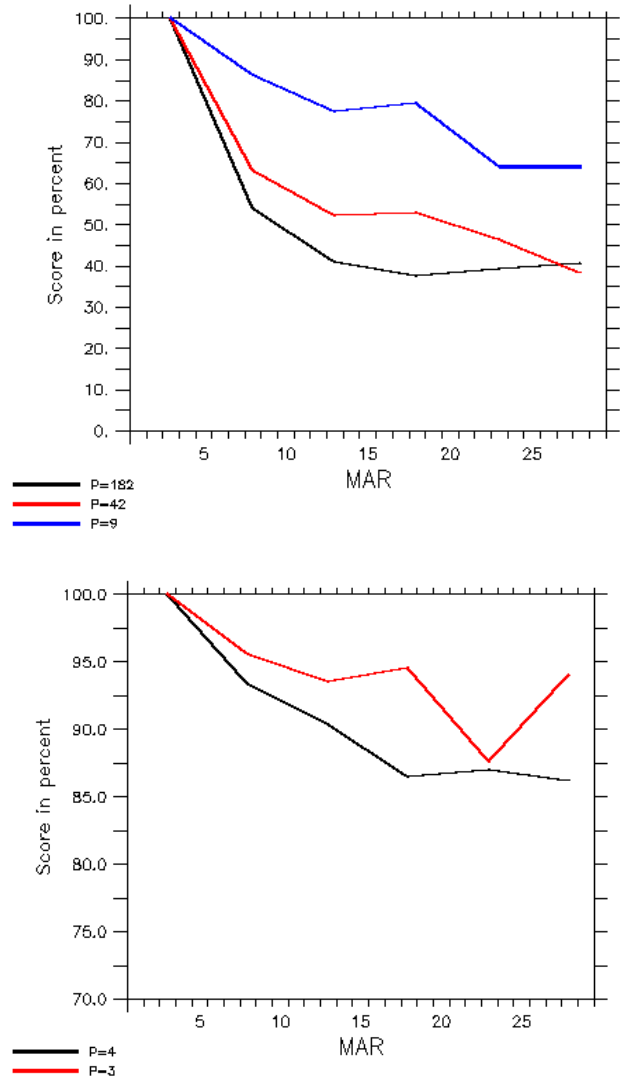


Fig. 9. Evolution of the adjustment error $E(t)$ at the parking depth computed from realistic surface data with different coverage: $P=182$, 42, 9 (a), and $P=3$, 4 (b).

results in Fig. 4 obtained with perfect datasets. With low data coverage (P equal to 9 and 42), the evolution of $E(t)$ is very close to that obtained with perfect datasets, with minimum values of 70% and 40% respectively obtained at the end of the assimilation period. In case of high data coverage (P equal to 182), on the other hand, the E decrease is quickly reduced after the third sequence. Its minimum is approximately 40%, while in the case of perfect data $E(t)$ reached 20%. This different behaviour can be explained considering that datasets of low coverage provide information of coarse resolution on the truth circulation, while shear errors have relatively small amplitudes (about 19% of the drift at parking depth) and would affect only fine scale information. In consequence, the obtained velocity estimates for low coverage

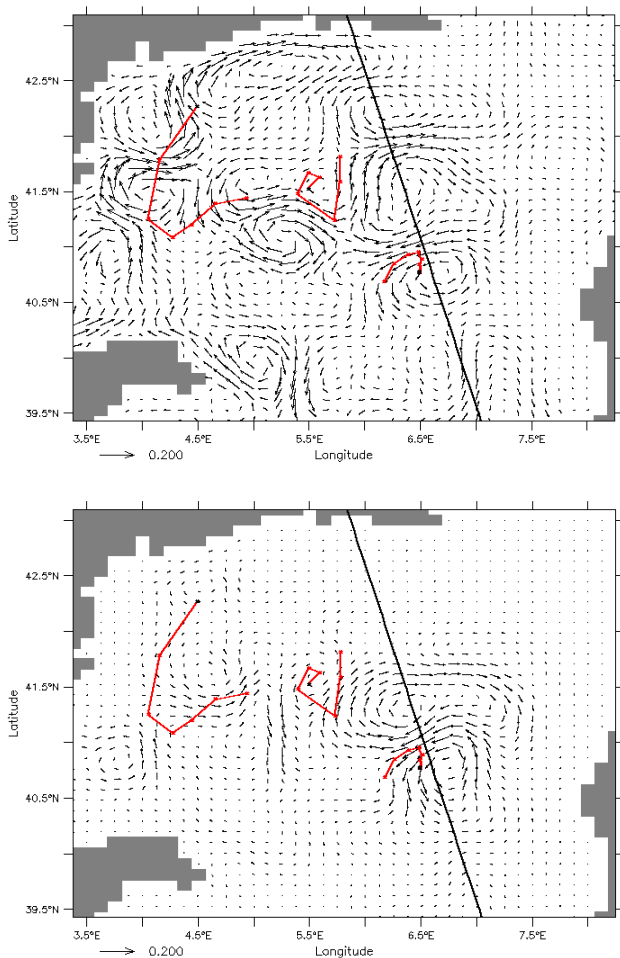


Fig. 10. Velocity misfits (in m/s) at the parking depth at the end of the assimilation period: truth minus background (a); estimate with $P=3$ minus background (b).

are likely to be less sensitive to the introduction of shear errors which only affect high resolution datasets.

A second characteristic of Argo floats concerns their typical spatial coverage. For example, there are typically three to six active floats in the North Western Mediterranean Sea in the framework of the MFSTEP project. So the simulated data sets have to be reduced to few trajectories. Corresponding adjustment errors $E(t)$ are represented in Fig. 9b for P equal to 3 and 4. $E(t)$ decreases in average indicating that the assimilation is effective, but the evolution appears non monotonous, especially for P equal to 3, with variations between sequences of the order of 5%. Minimum E values reached during the one month assimilation are between 85% and 95%.

Such a global assessment needs to be refined in order to understand what could be expected in the assimilation of real Argo data positions. In order to have a more detailed, even though qualitative assessment, we consider a horizontal map

at the parking depth of the truth minus background misfit ($\mathbf{u}^{\text{tru}} - \mathbf{u}^{\text{bck}}$) at the end of the integration and we compare it with the estimate minus background misfit ($\mathbf{u}^{\text{est}} - \mathbf{u}^{\text{bck}}$) obtained with P equal to 3. The corresponding fields and floats trajectories are shown in Fig. 10. A visual comparison shows that the assimilation tends to reproduce correctly the main structures, at least in the areas where data are available. The three floats sample different regions and have different impacts on the assimilation process. The most western float is located in a very energetic area. It is therefore characterized by important displacements. The other two floats are located in a more quiescent region, with meandering motion and reduced displacements. For the energetic float, the structures of velocity corrections obtained at each sequence appear to propagate quickly south-westward, as indicated by the structure situated north of the Balearic Islands. The propagation is likely to be due to a combination of advection processes and propagation of Rossby modes that are prominent in this area (Pinardi and Navarra, 1993). As a consequence, the velocity corrections which persist in the map (Fig. 10b) are mostly performed at the final sequences. Thus the global adjustment error $E(t)$ quickly reaches its threshold value. For the less energetic floats in the central basin, on the other hand, the local velocity correction appears less influenced by propagation. As the floats stay longer in the same zone, velocity estimates obtained at the end of the assimilation period result from the whole sequences of analysis. Particularly, the eastern trajectory infers a major correction on the background circulation to relocate northward the flow path. The trajectory located at the centre, which starts with slow displacements to finish with larger ones, leads to velocity corrections partially propagating westward where they contribute to the correction inferred by the previous float. Another part of the correction does not appear to move, and it leads to a local correction which is not in complete agreement with the truth minus background misfit at the end of the assimilation period.

6 Summary and concluding remarks

An assimilation method for Argo float positions has been developed and implemented. Investigations have been performed in a realistic model configuration of the Mediterranean Sea focusing on the North Western area. A first set of experiments have been performed to test the method in ideal conditions. Perfect data with no shear error are considered, corresponding to the initial and final positions of each cycle drift at the parking depth z_p . Relatively high coverages are considered, with a number of floats P ranging between 182 and 9. Two different procedures are tested: a “state of art” one, where for each sequence the analysis is performed at initial time and the model is re-run forward starting from the corrected initial conditions; and an “operational” one where

the analysis is performed at final time and the model is not re-run forward on the sequence.

The results show that the assimilation is effective for both procedures. The velocity field estimated at z_p converges toward the truth at increasing coverage. The global adjustment error E , computed over the region of interest, ranges between 20% for P equal to 182 and 70% for P equal to 9 after one month of assimilation. The correction projected in the water column is effective in a thick intermediate layer around z_p . For deeper layers or lower coverage (P equal to 9) the correction is lower but still significant, with values around $E=70\%$. Also the correction of the mass field TS appears significant and consistent, even for low coverage. The performance of the two procedures is similar, provided that the mass field is estimated as discussed in Sect. 2.2. Only when TS are not actively corrected (and let free to evolve and passively correct themselves), the “operational” procedure appears significantly less effective. Given that the “operational” procedure has a lower CPU cost and it is more compatible with standard procedures for the assimilation of other data sets, we consider it more convenient and we recommend it for future applications.

A second set of tests have been performed (with the “operational” procedure only) using data with realistic characteristics, with the goal of investigating the potential impact of Argo floats in observing systems. Realistic position data are considered, i.e. taken at the surface and therefore contaminated by shear errors. The results are compared with those for perfect data with P equal to 182, 42, 9. Accounting for shear errors shows a significant impact for high coverage, P equal to 182, increasing the final adjustment E from 20% to 40%. For lower coverage, the impact appears negligible, possibly because its fine scale effect is masked by the coarse resolution of the datasets used to describe the truth circulation. We have then considered the impact of realistically low coverage, considering P equal to 4, 3. The assimilation appears effective in the sense that the global adjustment error $E(t)$ tends to decrease in average, even though its final value is high, of the order of 85–95%. Locally, the assimilation effectively corrects the mesoscale structures sampled by the data. Even when the correction propagates dynamically (through advection and wave propagation) in regions with no data, it correctly modifies the model circulation with respect to the truth. Of course, as for any sparse dataset, the possibility of having spurious corrections in regions of data voids persists and it would be reinforced with longer experimental durations. Still after one month, this phenomenon appears very limited at least in the performed experiments, possibly because of the dynamically balanced nature of the correction.

In summary, the present results indicate that the assimilation of realistic Argo floats, despite shear error and sparseness, can provide significant corrections to model solutions and help converging toward the truth. These results are confirmed by a recent work by Taillandier et al. (2006b), where in-situ Argo data from the MFSTEP project in the Mediter-

anean Sea have been assimilated in the OPA model in the same region considered here. In the case of in-situ data, results cannot be tested against the truth which is obviously unknown. Nevertheless, the corrections appear significant and consistent with Argo float measurements and, at least qualitatively, also with independent current-meter measurements.

We notice that an intrinsic limit to the use of Argo data positions is the fact that the time interval Δt between successive observations is of the order of the typical Lagrangian time scale T_L . This implies that the details on the structure curvature are not contained in the data. In the present runs, we have not encountered any major problem related to this, differently from the previous work by Taillandier et al. (2006a) where floats in a strong vortex appeared to lead to spurious velocity corrections. This is probably due to the fact that in Taillandier et al. (2006a) the velocity field was reconstructed from float data alone, without assimilating them in a model, while here accounting for model dynamics helps providing additional information on the structures. On the other hand, even though no catastrophic event occurs in the present experiments, the float correction is in general less significant for structures of high curvature, such as small and intense vortices, with respect to circulation features with broader scales, such as major currents or meanders. This can be seen also in the assimilation of in-situ data (Taillandier et al., 2006b). When the floats enter strong currents and jets with long time and space correlations, the velocity correction is strong and consistent, while when floats loop in small structures such as submesoscale vortices (Testor and Gascard, 2003; Testor et al., 2005), the correction is weaker.

As a final remark we notice that, even though the present results are very positive, some aspects persist that can be further investigated and improved in future works. First of all, we notice that the results strongly depend on the statistical correlations used for projection of the velocity correction in the water column. While in the North Western Mediterranean Sea considered here the velocity correlation profiles appear well represented, it can be expected that in regions with higher variability and baroclinicity the approach might be less successful. Also, the use of the geostrophic balance in the mass correction limits the validity of the method to mesoscale open ocean or shelf flows. In principle, a powerful alternative to the statistical projection and simplified dynamics used here is a constraint based on the full model dynamics (e.g., Nodet, 2006). The counterpart is that the number of variables and parameters to be fitted to the data is much higher. So, for a reduced data set as for Argo floats, the choice of the most effective method is expected to be a compromise between the complexity of the considered dynamics and the type of available information. Finally, Argo floats measure also vertical profiles of temperature and salinity and the simultaneous assimilation of such TS profiles and positions would be a promising avenue for future works.

Acknowledgements. This work was supported by the European Commission (V Framework Program – Energy, Environment and Sustainable Development) as part of the MFSTEP project (contract number EVK3-CT-2002-00075) and by the Office of Naval Research grant N00014-05-1-0094. We wish to gratefully thank A. Bozec, K. Béranger, L. Mortier for their assistance, P. De Mey, A. Molcard, T. Özgökmen for useful discussions. Thanks to the MERCATOR project which provided the numerical model and the forcing from ECMWF.

Edited by: D. Webb

References

- Béranger, K., Mortier, L., and Crépon, M.: Seasonal variability of transports through the Gibraltar, Sicily and Corsica straits from a high resolution Mediterranean model, *Prog. Oceanogr.*, 66, 341–364, 2005.
- Gilbert, J.-C. and Lemaréchal, C.: Some numerical experiments with variable-storage quasi-Newton algorithms, *Math. Program.*, 45, 407–435, 1989.
- Ide, K., Kuznetsov, L., and Jones, C. K. R. T.: Lagrangian data assimilation for point vortex systems, *J. Turbul.*, 3, 53–59, 2002.
- Ishikawa, Y. I., Awaji, T., and Akimoto, K.: Successive correction of the mean sea surface height by the simultaneous assimilation of drifting buoys and altimetric data, *J. Phys. Oceanogr.*, 26, 2381–2397, 1996.
- Jackett, D. R. and McDougall, T. J.: Minimal adjustment of hydrographic profiles to achieve static stability, *J. Atmos. Ocean. Tech.*, 12, 381–389, 1995.
- Kamachi, M. and J. O'Brien: Continuous assimilation of drifting buoy trajectory into an equatorial Pacific Ocean model, *J. Mar. Syst.*, 6, 159–178, 1995.
- Kutsnetsov, L., Ide, K. and Jones, C. K. R. T.: A method for assimilation of Lagrangian data, *Mon. Weather Rev.*, 131, 2247–2260, 2003.
- Madec, G., Delecluse, P., Imbard, M., and Levy, C.: OPA8.1 ocean general circulation model reference manual, Technical Report LODYC/IPSL [http://www.lodyc.jussieu.fr/opa], 1998.
- Molcard, A., Griffa, A., and Özgökmen, T.: Lagrangian data assimilation in multilayer primitive equation models, *J. Atmos. Ocean. Tech.*, 22, 70–83, 2005.
- Molcard, A., Piterbarg, L. I., Griffa, A., Özgökmen, T. M., and Mariano A. J.: Assimilation of drifter positions for the reconstruction of the eulerian circulation field, *J. Geophys. Res.*, 108, 3056, 2003.
- Molcard, A., Poje, A. J., and Özgökmen, T. M.: Directed drifter launch strategies for Lagrangian data assimilation using hyperbolic trajectories, *Ocean Model.*, 12 (3–4), 268–289, 2006.
- Nodet, M.: Variational assimilation of Lagrangian data in oceanography, *Inverse Probl.*, 22, 245–263, 2006.
- Oschlies, A. and Willebrand, J.: Assimilation of Geosat altimeter data into an eddy-resolving primitive equation model of the North Atlantic Ocean, *J. Geophys. Res.*, 101, 14 175–14 190, 1996.
- Özgökmen, T. M., Molcard, A., Chin, T. M., Piterbarg, L. I., and Griffa, A.: Assimilation of drifter positions in primitive equation models of midlatitude ocean circulation, *J. Geophys. Res.*, 108, 3238, 2003.
- Park, J. J., Kim, K., King, B. A., and Riser, S. C.: An advanced method to estimate deep currents from profiling floats, *J. Atmos. Ocean. Tech.*, 22, 1294–1304, 2005.
- Pinardi, N. and Navarra, A.: Baroclinic wind adjustment processes in the Mediterranean Sea, *Deep Sea Research II*, 40, 1299–1326, 1993.
- Pinardi, N., Allen, I., Demirov, E., De Mey, P., Korres, G., Lascaratos, A., Le Traon, P.-Y., Maillard, C., Manzella, G., and Tziavos, C.: The Mediterranean ocean forecasting system: first phase of implementation (1998–2001), *Ann. Geophys.*, 21, 3–20, 2003.
- Poulain, P. M.: MEDARGO: A profiling float program in the Mediterranean, *Argonautics*, 6, 2, 2005.
- Raichich, F. and Rampazzo, A.: Observing system simulation experiments for the assessment of temperature sampling strategies in the Mediterranean Sea, *Ann. Geophys.*, 21, 151–165, 2003.
- Taillandier, V., Griffa, A., and Molcard, A.: A variational approach for the reconstruction of regional scale Eulerian velocity fields from Lagrangian data, *Ocean Model.*, 13, 1, 1–24, 2006a.
- Taillandier, V., Griffa, A., Poulain, P.-M., and Béranger, K.: Assimilation of Argo float positions in the North Western Mediterranean Sea and impact on ocean circulation simulations, *Geophys. Res. Lett.*, in press, 2006b.
- Talagrand, O.: A posteriori evaluation and verification of analysis and assimilation algorithms, *Proceedings of workshop on “Diagnostic of data assimilation systems”* ECMWF, 1999.
- Talagrand, O. and Courtier, P.: Variational assimilation of meteorological observations with the adjoint vorticity equation, *Q. J. Roy. Meteor. Soc.*, 113, 1311–1328, 1987.
- Testor, P. and Gascard, J.-C.: Large-scale spreading of deep waters in the Western Mediterranean Sea by submesoscale coherent eddies, *J. Phys. Oceanogr.*, 33, 75–87, 2003.
- Testor, P., Béranger, K., and Mortier, L.: Modelling the eddy field in the South Western Mediterranean: the life cycle of Sardinian eddies, *Geophys. Res. Lett.*, 32, L13602, 2005.
- Toner, M., Poje, A. J., Kirwan, A. D., Jones, C. K. R. T., Lipphardt, B. L., and Grosch, C. E.: Reconstructing basin-scale Eulerian velocity fields from simulated drifter data, *J. Phys. Oceanogr.*, 31, 1361–1376, 2001.